Report on FCC Computing Center

Patricia McBride

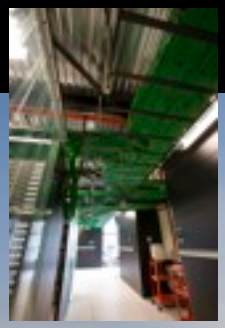
Deputy Head Computing Division

February 22, 2010

Recent Unscheduled Power Outages at FCC

- Message from February 17, 2010:
  - An unscheduled power outage occurred at the Feynman Computing Center (FCC) today at approximately 3:48AM. UPS1 and cooling were affected. The outage impacted the computing equipment in FCC computer rooms. There is no future outage scheduled at this time. Check the Service Desk web site for additional details.
- Message from February 9, 2010:
  - An unscheduled power outage occurred at the Feynman Computing Center (FCC) yesterday at approximately 12:25PM. The outage impacted the computing equipment in FCC computer rooms. The primary UPS electrical power and all cooling within the computer rooms was shutdown by a fault in the Emergency Power Off (EPO) circuit. The EPO is a safety mechanism intended to power down the computer rooms in case of an emergency in order to protect personnel and facilities. The Feynman building infrastructure was generally stable and uninterrupted during the outage within the computer rooms.
- http://computing.fnal.gov/xms/Services/Service_Desk

Fermilab Computing Facilities



Feynman Computing Center (FCC)

High Availability Services

Networking, Computer Security, BSS, Email, Web, Databases, Core Site Services

Tape Robotic Storage

UPS & Standby Power Generator

NEW HA capacity available in FY11

GRID Computing Center (GCC)

High Density Computational Computing

CMS, RUNII, GP Batch Worker nodes

Lattice HPC nodes

Tape Robotic Storage

UPS & taps for portable generators

Lattice Computing Center (LCC)

High Performance Computing (HPC)

Accelerator Sim, Cosmology nodes

No UPS; taps for portable generators

Incident

- We have four independent UPS circuits in FCC:
  - UPS-1 (576kVA)- many core services + BlueArc and other disk
  - UPS-2 (60kVA) - network services
  - UPS-3 (100kVA) - CMS services
  - UPS-4 (100kVA) - CDF services
- Due to every increasing demand for power in our computer rooms, we have been operating the UPS1 circuit at the highest total load since the power system was upgraded in 2001.
- All computers and cooling systems connected to UPS-1 were impacted when a breaker rated at 1400 amps per phase tripped. Computers on UPS2 (60kVA) , UPS3 (100kVA) and UPS4 (100kVA) were powered with no cooling for several hours.

Critical incident response

Lessens learned from the first outage was put to use for the 2nd incident.

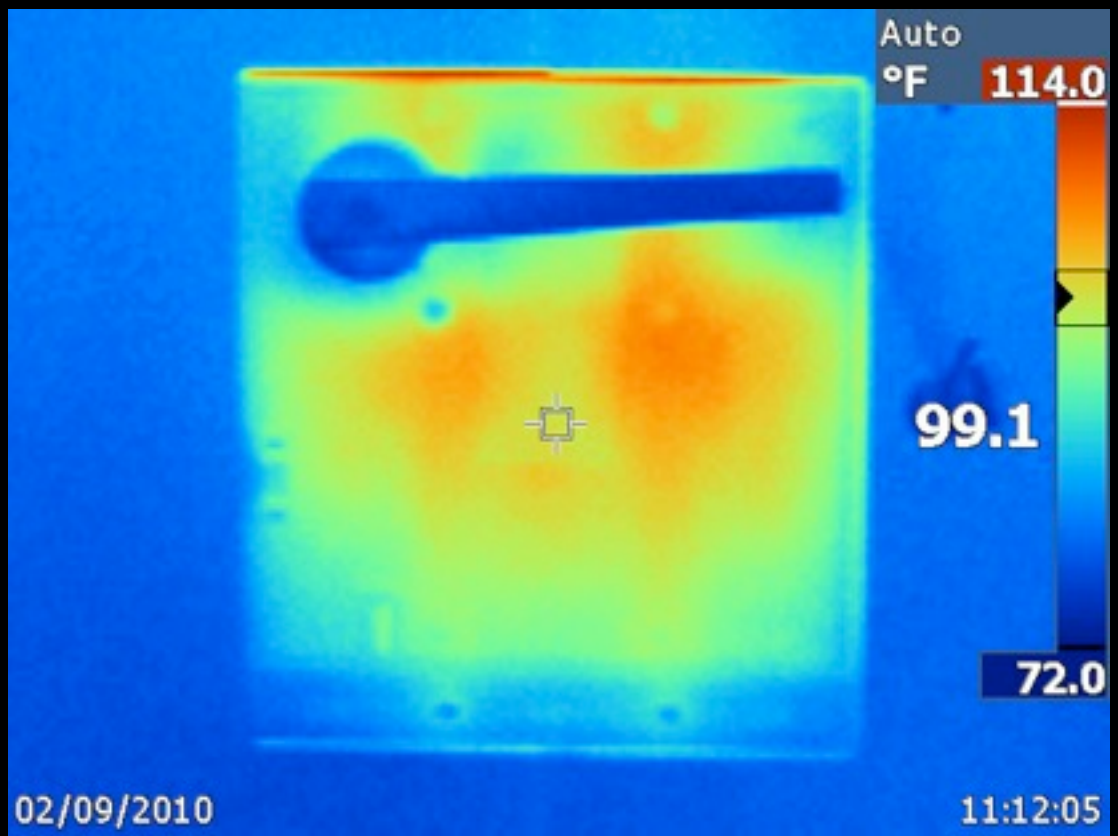
Eileen Berman served as the critical incident manager and handled coordination across the division and across the laboratory.

The division facilities team and FESS worked to discover the root cause of the 2 incidents.

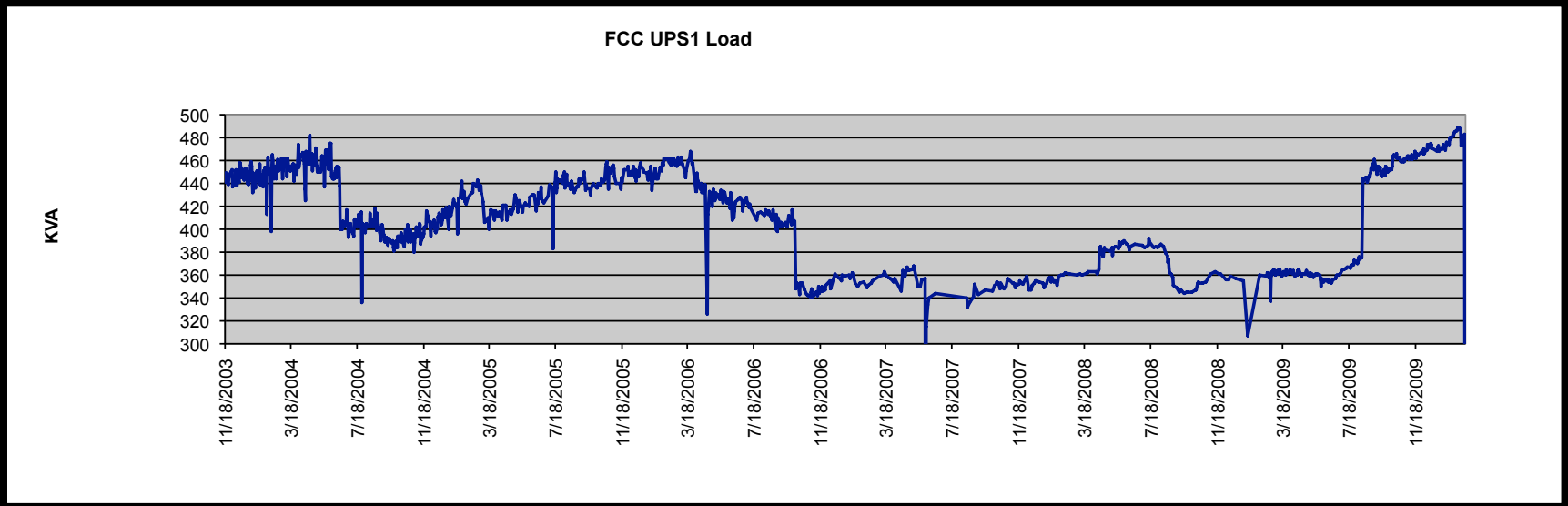
CD brought systems back online by 7pm Wednesday - but with power limit was reduced to 1120amps/phase.

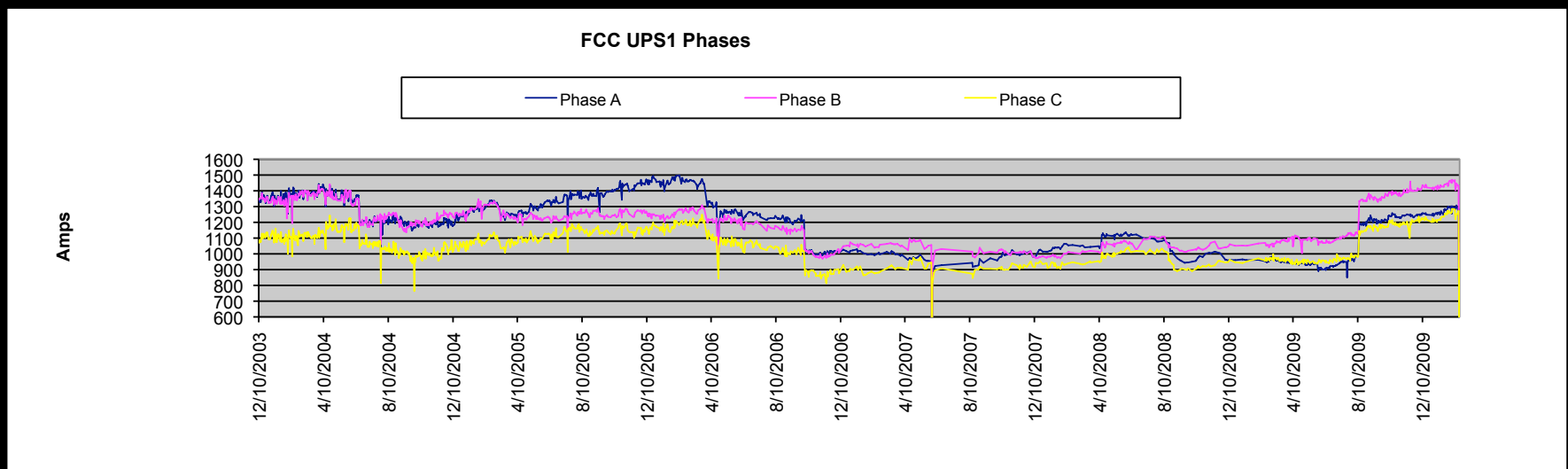
Breaker - Photos

Photo of the breaker taken one hour before the February 9 outage.



History UPS - 1





Planning for recovery

The Computing Division incident management team (including many people from across the division) has tried to reduce the impact of this outage and continued degradation of service. The CD team contacted the experiments, the divisions and sections to understand how to best restore essential services and to minimize the disruption.

Planning for recovery:

CD is working with FESS to exploring ways to increase the available power in a safe manner.

Alternatively, we are also exploring plans to move computing equipment out of FCC to allow everything to turn back on with a minimal amount of effort. Preliminary estimate is that we may need to move about 10 racks. The CD facilities group is evaluating if there is room (space, power, cooling, networking) in GCC for the relocated equipment.

It would be useful if we have capacity for new equipment that will be purchased in the next few months - before the completion of the new FCC3 HA computer facility.

Conclusions

The partial restoration of service is not a sustainable solution. The Computing Division and FESS are currently performing a root cause analysis of the breaker trip.

Until this analysis is complete, and to lessen the risk of another trip, the facilities team has recommended operating the breaker at a reduced load consistent with our understanding of its ability to run without further trips.

This means that we cannot now power up all services, and we may have to move some equipment to other CD computing rooms.

The Computing Division together with FESS are making an evaluation of the maximum power available and a plan for powering up all equipment as quickly as possible.

If you have questions about the availability or priority of systems or services, please contact the Fermilab Service Desk, x2345 or http://servicedesk.fnal.gov.

Thank you

- Thank you to the critical incident response team headed by Eileen Berman during the outage. The whole team stayed calm and communicated well.
- I would like to thank the CD facilities team for their dedicated support during this period.
- CD also wants to thank FESS for their help in evaluating the facilities infrastructure.
- And a thank you to all the users and staff for their patience and cooperation.

CD is working hard to restore service and will provide many updates during the coming days.